



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Philipp, Mario T.
- (ii) TITLE OF INVENTION: Surface Antigens and Proteins Useful in Compositions for the Diagnosis and Prevention of Lyme Disease
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Howson and Howson
  - (B) STREET: 501 Office Center Drive, Suite 210
  - (C) CITY: Fort Washington
  - (D) STATE: Pennsylvania
  - (E) COUNTRY: USA
  - (F) ZIP: 19034
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 10/632,780
  - (B) FILING DATE: 01-AUG-2003
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 09/445,803
  - (B) FILING DATE: 13-DEC-1999
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: PCT/US98/13551
  - (B) FILING DATE: 29-JUN-1998
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 60/051,271
  - (B) FILING DATE: 30-JUN-1997
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Bak, Mary E.
  - (B) REGISTRATION NUMBER: 31,215
  - (C) REFERENCE/DOCKET NUMBER: TUL2BUSA
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 215-540-9200
  - (B) TELEFAX: 215-540-5818

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1047 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: CDNA

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1047

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAG Lys 1	AAT Asn	AAT Asn	GAT Asp	CAT His 5	GAT Asp	AAT Asn	CAT His	AAG Lys	GGG Gly 10	ACT Thr	GTT Val	AAG Lys	AAT Asn	GCT Ala 15	GTT Val	48
GAT Asp	ATG Met	GCA Ala	AAG Lys 20	GCC Ala	GCT Ala	GAG Glu	GAA Glu	GCT Ala 25	GCA Ala	AGT Ser	GCT Ala	GCA Ala	AGT Ser 30	GCT Ala	GCT Ala	96
ACT Thr	GGT Gly	AAT Asn 35	GCA Ala	GCG Ala	ATT Ile	GGG Gly	GAT Asp 40	GTT Val	GTT Val	AAG Lys	AAT Asn	AGT Ser 45	GGG Gly	GCA Ala	GCA Ala	144
GCA Ala	AAA Lys 50	GGT Gly	GGT Gly	GAG Glu	GCG Ala	GCG Ala 55	AGT Ser	GTT Val	AAT Asn	GGG Gly	ATT Ile 60	GCT Ala	AAG Lys	GGG Gly	ATA Ile	192
AAG Lys 65	GGG Gly	ATT Ile	GTT Val	GAT Asp 70	GCT Ala	GCT Ala	GGA Gly	AAG Lys	GCT Ala	GAT Asp 75	GCG Ala	AAG Lys	GAA Glu	GGG Gly	AAG Lys 80	240
TTG Leu	GAT Asp	GCT Ala	ACT Thr	GGT Gly 85	GCT Ala	GAG Glu	GGT Gly	ACG Thr	ACT Thr 90	AAC Asn	GTG Val	AAT Asn	GCT Ala	GGG Gly 95	AAG Lys	288
TTG Leu	TTT Phe	GTG Val	AAG Lys 100	AGG Arg	GCG Ala	GCT Ala	GAT Asp	GAT Asp 105	GGT Gly	GGT Gly	GAT Asp	GCA Ala	GAT Asp 110	GAT Asp	GCT Ala	336
GGG Gly	AAG Lys	GCT Ala 115	GCT Ala	GCT Ala	GCG Ala	GTT Val	GCT Ala 120	GCA Ala	AGT Ser	GCT Ala	GCT Ala	ACT Thr 125	GGT Gly	AAT Asn	GCA Ala	384
GCG Ala 130	ATT Ile	GGA Gly	GAT Asp	GTT Val	GTT Val	AAT Asn 135	GGT Gly	GAT Asp	GTG Val	GCA Ala	AAA Lys 140	GCA Ala	AAA Lys	GGT Gly	GGT Gly	432
GAT Asp 145	GCG Ala	GCG Ala	AGT Ser	GTT Val	AAT Asn 150	GGG Gly	ATT Ile	GCT Ala	AAG Lys	GGT Gly 155	ATA Ile	AAG Lys	GGG Gly	ATT Ile	GTT Val 160	480
GAT Asp	GCT Ala	GCT Ala	GAG Glu	AAG Lys 165	GCT Ala	GAT Asp	GCG Ala	AAG Lys	GAA Glu 170	GGG Gly	AAG Lys	TTG Leu	AAT Asn	GCT Ala 175	GCT Ala	528
GGT Gly	GCT Ala	GAG Glu	GGT Gly 180	ACG Thr	ACT Thr	AAC Asn	GCG Ala	GAT Asp 185	GCT Ala	GGG Gly	AAG Lys	TTG Leu	TTT Phe 190	GTG Val	AAG Lys	576
AAT Asn	GCT Ala	GGT Gly 195	AAT Asn	GTG Val	GGT Gly	GGT Gly	GAA Glu 200	GCA Ala	GGT Gly	GAT Asp	GCT Ala	GGG Gly 205	AAG Lys	GCT Ala	GCT Ala	624
GCT Ala 210	GCG Ala	GTT Val	GCT Ala	GCT Ala	GTT Val	AGT Ser 215	GGG Gly	GAG Glu	CAG Gln	ATA Ile	TTA Leu 220	AAA Lys	GCG Ala	ATT Ile	GTT Val	672
CAT His	GCT Ala	GCT Ala	AAG Lys	GAT Asp	GGT Gly	GGT Gly	GAG Glu	AAG Lys	CAG Gln	GGT Gly	AAG Lys	AAG Lys	GCT Ala	GCG Ala	GAT Asp	720

225	230					235					240					
CGT Arg	ACA Thr	AAT Asn	CCC Pro	ATT Ile 245	GAC Asp	GCG Ala	GCT Ala	ATT Ile	GGG Gly 250	GGT Gly	GCG Ala	GGT Gly	GAT Asp	AAT Asn 255	GAT Asp	768
GCT Ala	GCT Ala	GCG Ala	GCG Ala 260	TTT Phe	GCT Ala	ACT Thr	ATG Met	AAG Lys 265	AAG Lys	GAT Asp	GAT Asp	CAG Gln	ATT Ile 270	GCT Ala	GCT Ala	816
GCT Ala	ATG Met	GTT Val 275	CTG Leu	AGG Arg	GGA Gly	ATG Met	GCT Ala 280	AAG Lys	GAT Asp	GGG Gly	CAA Gln	TTT Phe 285	GCT Ala	TTG Leu	AAG Lys	864
GAT Asp	GCT Ala 290	GCT Ala	GCT Ala	GCT Ala	CAT His	GAA Glu 295	GGG Gly	ACT Thr	GTT Val	AAG Lys	AAT Asn 300	GCT Ala	GTT Val	GAT Asp	ATA Ile	912
ATA Ile 305	AAG Lys	GCT Ala	GCT Ala	GCG Ala	GAA Glu 310	GCT Ala	GCA Ala	AGT Ser	GCT Ala	GCA Ala 315	AGT Ser	GCT Ala	GCT Ala	ACT Thr	GGT Gly 320	960
AGT Ser	GCA Ala	GCA Ala	ATT Ile	GGG Gly 325	GAT Asp	GTT Val	GTT Val	AAT Asn	GGT Gly 330	AAT Asn	GGA Gly	GCA Ala	ACA Thr	GCA Ala 335	AAA Lys	1008
GGT Gly	GGT Gly	GAT Asp	GCG Ala 340	AAG Lys	AGT Ser	GTT Val	AAT Asn	GGC Gly 345	ATT Ile	GCT Ala	AAG Lys	GGA Gly				1047

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 349 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Lys 1	Asn	Asn	Asp	His 5	Asp	Asn	His	Lys	Gly 10	Thr	Val	Lys	Asn	Ala 15	Val
Asp	Met	Ala	Lys 20	Ala	Ala	Glu	Glu 25	Ala	Ala	Ser	Ala	Ala	Ser 30	Ala	Ala
Thr	Gly	Asn 35	Ala	Ala	Ile	Gly	Asp 40	Val	Val	Lys	Asn	Ser 45	Gly	Ala	Ala
Ala	Lys 50	Gly	Gly	Glu	Ala	Ala 55	Ser	Val	Asn	Gly	Ile 60	Ala	Lys	Gly	Ile
Lys 65	Gly	Ile	Val	Asp	Ala 70	Ala	Gly	Lys	Ala	Asp 75	Ala	Lys	Glu	Gly	Lys 80
Leu	Asp	Ala	Thr	Gly 85	Ala	Glu	Gly	Thr	Thr 90	Asn	Val	Asn	Ala	Gly 95	Lys
Leu	Phe	Val	Lys 100	Arg	Ala	Ala	Asp	Asp 105	Gly	Gly	Asp	Ala	Asp 110	Asp	Ala
Gly	Lys 115	Ala	Ala	Ala	Ala	Val	Ala 120	Ala	Ser	Ala	Ala	Thr 125	Gly	Asn	Ala

Ala Ile Gly Asp Val Val Asn Gly Asp Val Ala Lys Ala Lys Gly Gly  
130 135 140  
Asp Ala Ala Ser Val Asn Gly Ile Ala Lys Gly Ile Lys Gly Ile Val  
145 150 155 160  
Asp Ala Ala Glu Lys Ala Asp Ala Lys Glu Gly Lys Leu Asn Ala Ala  
165 170 175  
Gly Ala Glu Gly Thr Thr Asn Ala Asp Ala Gly Lys Leu Phe Val Lys  
180 185 190  
Asn Ala Gly Asn Val Gly Gly Glu Ala Gly Asp Ala Gly Lys Ala Ala  
195 200 205  
Ala Ala Val Ala Ala Val Ser Gly Glu Gln Ile Leu Lys Ala Ile Val  
210 215 220  
His Ala Ala Lys Asp Gly Gly Glu Lys Gln Gly Lys Lys Ala Ala Asp  
225 230 235 240  
Arg Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ala Gly Asp Asn Asp  
245 250 255  
Ala Ala Ala Ala Phe Ala Thr Met Lys Lys Asp Asp Gln Ile Ala Ala  
260 265 270  
Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu Lys  
275 280 285  
Asp Ala Ala Ala Ala His Glu Gly Thr Val Lys Asn Ala Val Asp Ile  
290 295 300  
Ile Lys Ala Ala Ala Glu Ala Ala Ser Ala Ala Ser Ala Ala Thr Gly  
305 310 315 320  
Ser Ala Ala Ile Gly Asp Val Val Asn Gly Asn Gly Ala Thr Ala Lys  
325 330 335  
Gly Gly Asp Ala Lys Ser Val Asn Gly Ile Ala Lys Gly  
340 345

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 283 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCCGCTGGAT GGTGGTGAGA AGCAGGGTAA GAAGGCTGCG GATCGTACAA ATCCCATTGA	60
CGCGGCTATT GGGGGTGCGG GTGATAATGA TGCTGCTGCG GCGTTTGCTA CTATGAAGAA	120
GGATGATCAG ATTGCTGCTG CTATGGTTCT GAGGGGAATG GCTAAGGATG GGCAATTTGC	180
TTTGAAGGAT GCTGCTGCTG CTCATGAAGG GACTGTTAAG AATGCTGTTG ATATAATAAA	240

GGCTGCTGCG GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGT

283

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTATTATAT CAACAGATTC TTAACAGTCC CTTTCATGAGC AGCAGCAGCA TCCTTCAAAG	60
CAAATTGCCC ATCCTTAGCC ATTCCCCTCA GAACCATAGC AGCAGCAATC TGATCATCCT	120
TCTTCATAGT AGCAAACGCC GCAGCAGCAT CATTATCACC CGCACCCCCA ATAGCCGCGT	180
CAATCGGATT TGTACGATCC GCAGCCTTCT TACCCTGCTT CTCACCACCA TCC	233

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 194 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGTGCAAGC TGGGTTGAAG AAGGTTGGGG ATGTTGTAA GAATAGTGAG GCAAAAGATG	60
GTGATGCGGC GAGTGTTAAT GGGATTGCTA AGGGGATAAA GGGGATTGTT GATGCTGCTG	120
AGAAGGCTGA TGCGAAGGAA GGGAAGTTGG TATGTGGCTG GTGCTGCTGG TGAAACTAAC	180
AAGGAAGCGG CCGC	194

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGGCCGCTT GAGGAAGCTG CAAGTGCTGC AAGTGCTGCT ACTGGTAATG CAGCGATTGG	60
---	----

GGATGTTGTT AAGAATAGTG GGGCAGCAGC AAAAGGTGGT GAGGCGGCGA GTGTTAATGG	120
GATTGCTAAG GGGATAAAGG GGATTGTTGA TGCTGCTGGA AAGGCTGATG CGAAGGAAGG	180
GAAGTTGGAT GCTACTGGTG CTGAGGGTAC GACTAACGTG AATGCTGGGA AGTTGTTTGT	240
GAAGAGGGCG GCTGATGATG GTGGTGATGC AGATGATGCT GGGGAAGGCTG CTGCTGCGGT	300
TGCTGCAAGT GCTGCTACTG GTAATGCAGC GATTGGAGAT GTTGTTAATG GTGATGTGGC	360
AAAACAAAA	369

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 142 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGGATGGTG ATGATAAGCA GGGTAAGAAG GCTGAGGATG CTACAAATCC GATTGACGCG	60
GCTATTGGGG GTGCAGGTGC GGGTGCTAAT GCTGCTGCGG CGTTTAATAA TATGAAGAAG	120
GATGATCAGA TTGAGCGGCC GC	142

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 210 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGTGAACTA ACAAGGATGC TGGGAAGTTG TTTGTGAAGA AGAATGGTGA TGATGGTGGT	60
GATGCAGGTG ATGCTGGGAA GGCTGCTGCT GCGGTTGCTG CTGTTAGTGG GGAGCAGATA	120
TTAAAAGCGA TTGTTGATGC TGCTAAAGAT GGTGATAAGA CGGGGGTTAC TGATGTAAAG	180
GATGCTACAA ATCCGATTGA CGCGGCTATT	210

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 236 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATATAATAA AGGCTGCTGC GAAGCTGCAA GTGCTGCAAG TGCTGCTACT GGTAGTGCAG	60
CAATTGGGGA TGTGTGTTAAT GGTAATGGAG CAACAGCAAA AGGTGGTGAT GCGAAGTGTT	120
AATGGGATTG CTAAGGGGAT AAAGGGGATT GTTGATGCTG CTGAGAAGGC TGATGCGAAG	180
GAAGGGAAGT TGGATGTGGC TGGTGATGCT GGTGAAACTA ACAAGGAAGC GGCCGC	236

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 199 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGAGAGGAT CTCATCACCA TCACCATCAC ACGGATCCCC CGGGCTGCAG GAATTCGCGG	60
CCGCTGAAGG CTGATGCGAA GGAAGGGAAG TTGGATGTGG CTGGTGCTGC TGGTGAACT	120
AACAAGGATG CTGGGAAGTT GTTTGTGAAG AAGAATAATG AGGGTGGTGA AGCAAATGAT	180
GCTGGGAAGG CTGCTGCTG	199

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCCGCTGGAT GATCAGATTG CTGCTGCTAT GGTTGTGAGG GGAATGGCTA AGGATGGGCA	60
GTTTGCTTTG AAGGATGATG CTGCTAAGGA TGGAGATAAA ACGGGGGTTG CTGCGGATGT	120
GAAAATCCGA TTGACGCGGC TATTGGGGGT GCGGATGCTG ATGCTGCGGC GTTTAATAAG	180
GAGGGGATGA AGAAGGATGA TCAGATTGCT GCTGCTATGG TTCTGAGGGG AATGGCTAAG	240
GATGGGCAGT TTGCTTTGAC GAATAATGCT GC	272

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 289 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACTGTTAAGA ATGCTGTTGA TATAATAAAG GCTGCTGCGG AAGCTGCAAG TGCTGCAAGT	60
GCTGCTACTG GTAGTGACGC AATTGGGGAT GTTGTTAATG GTAATGGAGC AACAGCAAAA	120
GGTGGTGATG CGAAGAGTGT TAATGGGATT GCTAAGGGGA TAAAGGGGAT TGTTGATGCT	180
GCTGAGAAGG CTGATGCGAA GGAAGGGAAG TTGGATGTGG CTGGTGATGC TGGTGAAACT	240
AACAAGGATG CTGGGAAGTT GTTTGTGAAG AACAATGGTA ATGAGGGTA	289

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 142 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 2..142

- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION: 2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

G CCG CTT ACA AAT CCG ATT GAC GCG GCT ATT GGG GGG AGT GCG GAT	46
Pro Leu Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ser Ala Asp	
1 5 10 15	
CGT AAT GCT GAG GCG TTT GAT AAG ATG AAG AAG GAT GAT CAG ATT GCT	94
Arg Asn Ala Glu Ala Phe Asp Lys Met Lys Lys Asp Asp Gln Ile Ala	
20 25 30	
GCT GCT ATG GTT CTG AGG GGA ATG GCT AAG GAT GGG CAG TTT GCT TTG	142
Ala Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu	
35 40 45	

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 47 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear



(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro Leu Thr Asn Pro Ile Asp Ala Ala Ile Gly Gly Ser Ala Asp Arg  
1 5 10 15  
Asn Ala Glu Ala Phe Asp Lys Met Lys Lys Asp Asp Gln Ile Ala Ala  
20 25 30  
Ala Met Val Leu Arg Gly Met Ala Lys Asp Gly Gln Phe Ala Leu  
35 40 45